DEC 1 0 1987

GAO/NSIAD-85-68

BY THE COMPTROLLER GENERAL

Report To The Chairman, Senate Committee On Governmental Affairs

OF THE UNITED STATES

COMPLETED

Production Of Some Major Weapon Systems Began With Only Limited Operational Test And Evaluation Results

This report addresses the sufficiency of operational test and evaluation on concurrently developed and produced major weapon systems and the adequacy of reports on test results to the Congress.

Overall, GAO found that when major weapon systems are concurrently developed and produced, planned operational test and evaluation critical to assessing mission performance was typically not performed before beginning the production phase. GAO also found that available test results are sometimes omitted or misrepresented in reports to the Congress, creating misleading impressions of a weapon system's performance capabilities.

GAO makes recommendations to enhance DOD and congressional oversight on the performance of major weapon systems before production start-up.

UNCLASSIFIED

GAO/NSIAD-85-68
JUNE 19, 1985

Request for copies of GAO reports should be sent to:

U.S. General Accounting Office
Document Handling and Information
Services Facility
P.O. Box 6015
Gaithersburg, Md. 20877

Telephone (202) 275-6241

The first five copies of individual reports are free of charge. Additional copies of bound audit reports are $3.25 each. Additional copies of unbound report (i.e., letter reports) and most other publications are $1.00 each. There will be a 25% discount on all orders for 100 or more copies mailed to a single address. Sales orders must be prepaid on a cash, check, or money order basis. Check should be made out to the "Superintendent of Documents".

COMPTROLLER GENERAL OF THE UNITED STATES
WASHINGTON D.C. 20548

B-205620

The Honorable William V. Roth, Jr.
Chairman, Committee on Governmental
Affairs
United States Senate

Dear Mr. Chairman:

This report was prepared at your request and addresses the issue of whether sufficient operational test and evaluation results are obtained before production begins on major weapon systems that are concurrently developed and produced. It also describes the adequacy of operational test results reported in the Department of Defense's Congressional Data Sheets.

We are sending copies of this report to interested congressional committees; the Director, Office of Management and Budget; and to the Secretary of Defense.

Sincerely yours,

Charles A. Bowsher

Comptroller General
of the United States

Operational test and evaluation (OT&E) is an essential tool to use in deciding whether to begin production of a major weapon program costing billions of dollars. When major weapon programs do not undergo sufficient OT&E before production, the risk of costly redesign and modification after deployment increases. Risks can be increased further when weapon programs are expedited by producing certain parts of systems while other parts are being developed--a management practice known as concurrency. Concurrency can be an effective technique to expedite the development and production of weapon systems, provided it is well planned and controlled.

The Chairman of the Senate Committee on Governmental Affairs requested GAO to determine

--whether the Department of Defense (DOD) obtains sufficient OT&E results before production begins on major weapon systems that are concurrently developed and produced and

--the adequacy of OT&E information contained in DOD's annual Congressional Data Sheets.

BACKGROUND

GAO selected major weapon system programs (see p. iii) and assessed the provision made to perform OT&E before production start-up, the implementation of the planned OT&E, and the results achieved and reported from the OT&E that was performed.

Prior to fielding a system for combat use, test and evaluation is the primary means used to identify risks and demonstrate that a system has met performance requirements and can be advanced to the next acquisition phase.

Two major types of weapon systems' test and evaluation are performed in DOD and serve distinctly different purposes. Developmental test and evaluation is designed to assist the engineering design and development process and to verify that specifications and objectives are met. Developmental test and evaluation is normally accomplished or managed by the agency responsible for developing the weapon system.

In contrast, OT&E assesses a weapon system's effectiveness and suitability when operated and maintained by military personnel in a realistic combat environment and against a simulated enemy

threat. OT&E is conducted by the military departments' operational test agencies, that are intended to be separate and distinct from a system's developing, procuring, and using commands.

DOD permits the military services to "accept system performance growth after deployment" when the urgency to counter a threat transcends high technical, cost, and supportability risk. Nevertheless, OT&E results are required to be available before making major decisions regarding production start-up--even though the necessary corrective actions identified through OT&E may be postponed until after production begins.

RESULTS IN BRIEF

In all of the concurrently developed and produced weapon systems GAO reviewed, DOD did not obtain OT&E results critical to assessing mission performance before production start-up--even though DOD had initially planned to have these tests results available before making such decisions. GAO recognizes that concurrently developed and produced weapon system programs may be justified on an exception basis to meet certain extreme exigencies. However, concurrent weapon system programs should not prevent decisionmakers from having information to assess weapon system operational performance before production start-up.

In addition, GAO reviewed the OT&E results reported to the Congress in Congressional Data Sheets. These documents did not adequately disclose weapon system performance risks and shortcomings, critical issues, and other key conclusions obtained from the tests and evaluations that were conducted. One reason this situation exists may be that the military services' instructions on preparing Data Sheets do not contain specific guidance on the type of information that should be included in these documents.

PRINCIPAL FINDINGS

Because significant portions of planned OT&E were only partially completed on all of the five systems GAO reviewed, those systems began production without having adequately demonstrated whether or not performance requirements were met in a representative operational environment. GAO believes there is a need to independently assess the risks associated with such decisions. For example, at the time production began on the F/A-18 aircraft, it had not undergone critical

aspects of planned OT&E that was called for in the program schedule because of the system's immature development at that time. GAO points out that expensive retrofits were required on F/A-18 production models to correct problems identified during operational testing performed after the production decision was made.

The problem of insufficient OT&E before production start-up prevailed with all the other major weapon systems included in GAO's review--High Speed Antiradiation Missile, Air Launched Cruise Missile, B-1B Bomber, and the Sergeant York Air Defense Gun. For example, the Sergeant York program had little performance information to measure its reliability and maintainability before production began. OT&E results needed to assess critical aspects of a weapon systems performance are important because they provide early identification of problems and help prevent costly retrofits and performance shortcomings. (See pp. 10 through 21.)

GAO also found that operational test results omitted or misrepresented in Congressional Data Sheets created misleading impressions of a weapon system's demonstrated performance capabilities. For example, the Data Sheet for the Air Launched Cruise Missile did not adequately portray weapon system performance in a combat representative environment because launches were described as a partial success when in fact missiles had crashed. Further, the Data Sheet on the AEGIS class cruiser did not discuss certain problems that were cited in test reports such as the need to improve computer program performance which is critical for the system to be operationally effective. (See pp. 27 through 32 for additional examples.)

RECOMMENDATIONS

The Secretary of Defense should:

--Improve the quality of information to decisionmakers by assuring, through DOD's Director of OT&E, that test results critical to assessing mission performance are available before production start-up.

--Require that the Office of the Director of OT&E provide its views to the Senate and House Committees on Appropriations and Armed Services on the impact of not performing OT&E critical to assessing mission performance before production begins. This is particularly important when planned OT&E is not accomplished.

--Improve the quality of reporting OT&E results to the Congress by directing the military departments to expand their implementing instructions on preparing Congressional Data Sheets to include a more complete and accurate portrayal of OT&E results on major weapon programs.

AGENCY COMMENTS

DOD acknowledged that the programs GAO reviewed did not complete as much OT&E as is usually required before a weapon system enters production but noted that other factors such as cost and urgency persuaded DOD to proceed into production. DOD also stated that in almost all cases the production rate was limited and test results could be considered before authorization of full production. However, DOD agreed with GAO's recommendation that test results critical to assessing mission performance be available before production start-up, and stated that the Office of the Director of OT&E will be able to assure that in the future adequate testing is planned.

The draft of the GAO report proposed that when sufficient OT&E results are not available, the Congress should be given justification by the Director of OT&E concerning the necessity and risks of advancing systems into production. DOD disagreed with this proposal stating that its decisionmakers must decide when the risk of advancing a system into production is acceptable.

GAO agrees that the final decision rests with the Secretary of Defense and modified its recommendation. GAO's fundamental objective in making this recommendation is to assure that congressional decisionmakers are fully informed on the impact of starting production when sufficient OT&E results are not available.

DOD agreed with GAO's recommendation for improving the quality of reporting OT&E results to the Congress.

Contents

		Page
EXECUTIVE SUMMARY		i
CHAPTER		
1	INTRODUCTION	1
	DOD's policy on concurrency	1
	Acquisition strategy and program structure	2
	Test and evaluation	2
	Reporting OT&E results to the Congress	4
	Objectives, scope, and methodology	6
2	CONCURRENT DEVELOPMENT AND PRODUCTION WITHOUT SUFFICIENT OT&E LIMITS ABILITY TO PREDICT WEAPON PERFORMANCE	9
	Slippage of scheduled tests without concomitant delays in Navy production programs	10
	The Army's Sergeant York weapon program began production without the benefit of sufficient OT&E results	15
	The Air Force's ALCM and B-1B weapon programs began production before OT&E results were available	17
	Conclusions	21
	Recommendations	22
	Agency comments and our evaluation	22
3	THE QUALITY OF CONGRESSIONAL DATA SHEETS NEED FURTHER IMPROVEMENT TO BE MORE USEFUL TO THE CONGRESS	24
	Preparing and reviewing OT&E sections of Congressional Data Sheets	24
	Data Sheets for selected Navy systems omit important test results and information on testing limitations	27
	Potential effects of known deficiencies omitted from three Air Force Data Sheets	29
	Data Sheets for selected Army systems do not always explain significance of test limitations	31
	Conclusions	32
	Recommendation	33
	Agency comments and our evaluation	33

APPENDIX

		Page
I	Description of selected weapon systems	34
II	Letter from the Acting Director, Office of Operational Test and Evaluation dated February 12, 1985, commenting on our draft report	36

ILLUSTRATIONS

Examples of program structure	3
Five concurrent weapon programs	10

ABBREVIATIONS

AFOTEC	U.S. Air Force Operational Test and Evaluation Center
ALCM	Air Launch Cruise Missile
DOD	Department of Defense
HARM	High Speed Antiradiation Missile
OPTEVFOR	U.S. Navy Operational Test & Evaluation Force
OSD	Office of the Secretary of Defense
OT&E	operational test and evaluation
OTEA	Operational Test and Evaluation Agency

CHAPTER 1

INTRODUCTION

In a request by the Chairman of the Senate Committee on Governmental Affairs, we were asked to examine

--whether the Department of Defense (DOD) will obtain sufficient operational test and evaluation (OT&E)[1] results before production begins on major weapon systems that are concurrently developed and produced and

--the adequacy of OT&E information reported in DOD's annual Congressional Data Sheets.

The Chairman indicated that several witnesses in congressional testimony have stated that

> ". . . Often, the reduced time allotted for operational testing appears to have forced test agencies to rely heavily on developmental testing for operational test data. In addition, reduced testing times for operational tests may be resulting in inadequate testing and may encourage DOD officials to rely on incomplete test results to make production decisions on major weapon systems."

In this report, we discuss (1) DOD's procedures and practices for concurrent development and production of major weapon systems focusing on the adequacy and timeliness of OT&E results used to justify production start-up[2] decisions and (2) the quality of OT&E information contained in Congressional Data Sheets.

DOD'S POLICY ON CONCURRENCY

DOD's policy on major weapon system acquisition stresses the importance of minimizing the time to develop, produce, and deploy major systems for use by operational forces. To accomplish this objective, DOD's policy permits the services to build

---

[1]OT&E addresses how well the system can be expected to perform in the operational or combat environment, how it should be employed, and whether the system can be operated and maintained effectively by military personnel.

[2]In this report, we will use the term, production start-up, to mean the beginning of the production phase of a weapon program. Various terms have been used to describe production start-up, such as limited production, low-rate initial production, and pilot production. However, all these terms equate to the start-up of the production line. This report focuses on OT&E performed on a weapon program before production start-up.

concurrency into their weapon program structure. In this report, concurrency is defined as the overlap in time between the development of a weapon system and its production. In a nonconcurrent program, development is usually completed before production begins. In a concurrent program, production is started while development is still underway.

DOD's policy provides a framework for applying concurrency and requires documentation that would substantiate the need for concurrency. DOD's policy also specifies that the degree of concurrency will be based on the savings in acquisition time balanced against cost, risk, and urgency of the mission need in each acquisition program.

A weapon system program with planned concurrency should allow for special attention to OT&E so that performance risks resulting from a shortened acquisition time do not affect the planned deployment date. The decision whether to employ concurrency is to be part of the acquisition strategy that is planned at the inception of each major system acquisition.

ACQUISITION STRATEGY AND PROGRAM STRUCTURE

An acquisition strategy is the overall plan for executing a major weapon system program. The strategy should set forth such things as the objectives, resources, management assumptions, and program structure.

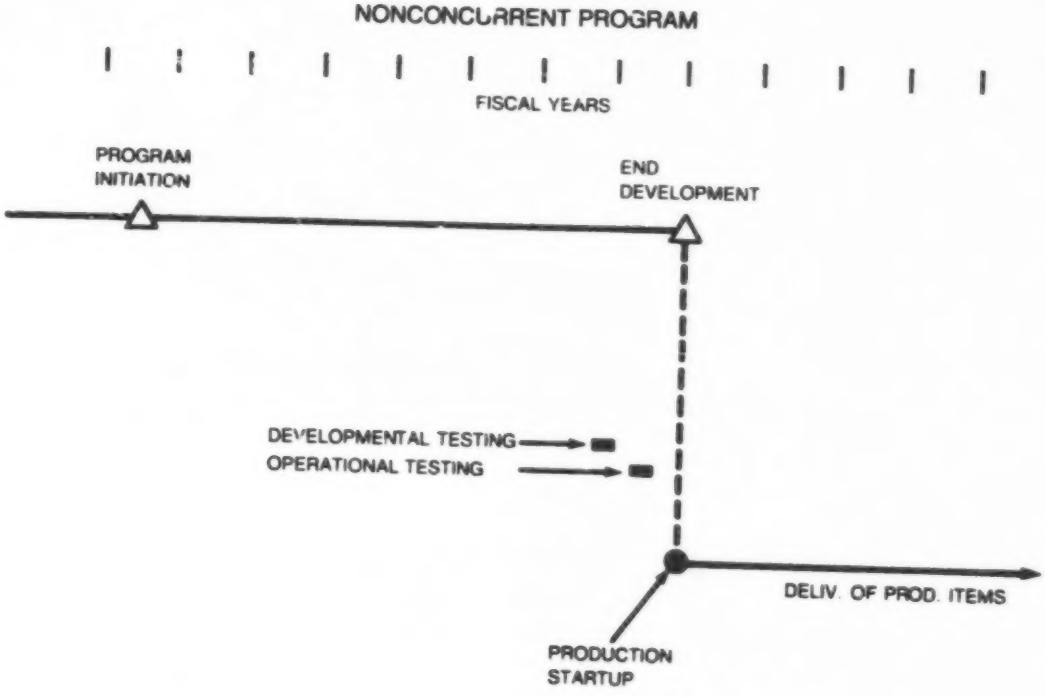
The program structure should identify such things as the proposed weapon system development phases, test and evaluation periods, and planned concurrency. The program structure is a very important aspect of DOD's acquisition strategy. Illustrated on the next page are two simplified schematics of program structure, one a nonconcurrent and the other a concurrent program.

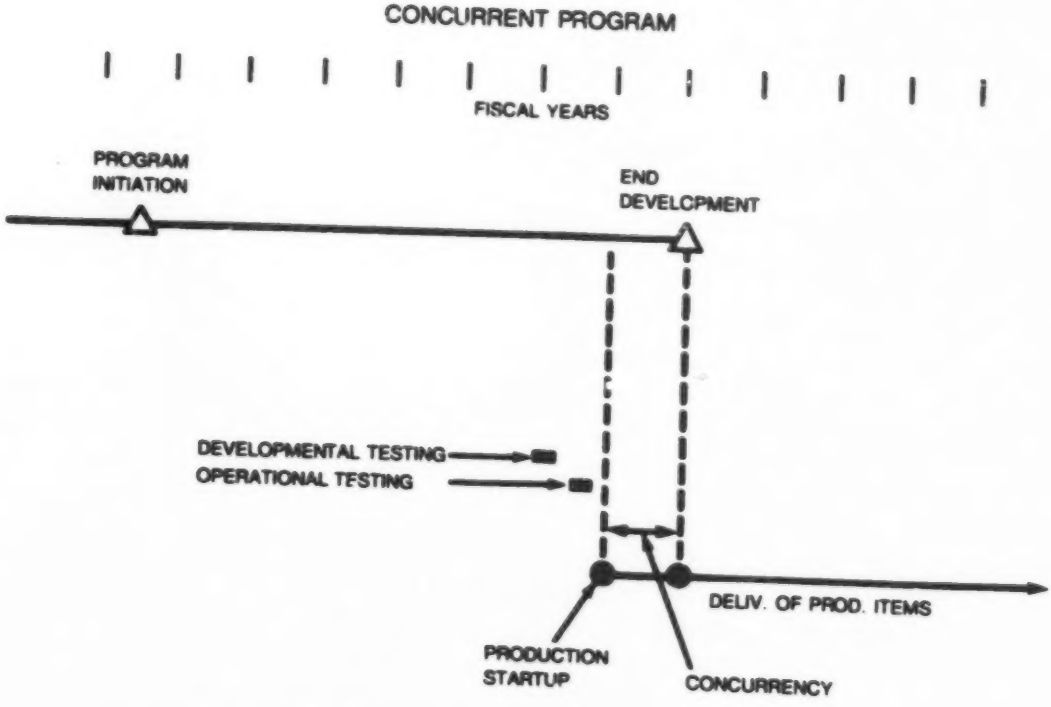
TEST AND EVALUATION

Short of war, test and evaluation is the primary means of assessing weapon system performance. Prior to fielding a system for combat use, the role of test and evaluation is to demonstrate that a system has met performance requirements and warrants advancing to the next phase.

Two types of weapon system's test and evaluation serve distinctly different purposes. Developmental test and evaluation is designed to assist the engineering design and development process and to verify that technical specifications are met. Developmental test and evaluation is normally accomplished or managed by the agency responsible for developing the weapon system.

EXAMPLES OF PROGRAM STRUCTURE





In contrast, OT&E is designed to assess system operational effectiveness and suitability,[3] when operated and maintained by military personnel, in a realistic combat environment and against a simulated enemy threat. A military service agency, separate and distinct from the developing/procuring/using agency, is responsible for managing and conducting operational testing and reporting test results to DOD.

Developmental and operational testing may be combined when significant cost and time benefits will result, provided that the necessary resources, test conditions, and test data can be obtained.

DOD's policy on OT&E

DOD's policy states that successful accomplishment of test and evaluation objectives will be a key requirement to support the level of commitment required at each key decision point. For example, production start-up, an important milestone, must be supported by sufficient OT&E[4] to estimate operational effectiveness and suitability, including logistics supportability.

REPORTING OT&E RESULTS TO THE CONGRESS

Adequate and timely OT&E can provide valuable information for the Congress to assess a weapon system's performance prior to production. Legislation was enacted in 1971 requiring DOD, among other things, to report OT&E results in annual reports (Congressional Data Sheets) to the Congress. Public Law 92-156 states that annual reports should include the results of all OT&E that has been conducted, or if OT&E has not been conducted, a statement of the reasons therefore and the results of such other test and evaluation as has been conducted.

---

[3]Operational effectiveness is the capability of a system to accomplish its mission. Operational suitability is the system's capability to be satisfactorily placed in field use considering, among other things, the capability to operate, maintain, and support the system.

[4]DOD's policy on test and evaluation states that the long design, engineering, and construction period of a major ship will "normally" preclude completion of tests before advancement from one acquisition phase to another. In addition, DOD's policy permits nuclear subsystems to be excluded from certain test and evaluation procedures and practices.

Under the law, DOD is required to provide the annual reports to the Committees on Armed Services and Appropriations on OT&E results for major weapon systems for which procurement funds were requested. The reports have been provided each year since 1973. These test results and other information are included in Congressional Data Sheets submitted annually to the the President submits the budget. The quality of OT&E information contained in the Data Sheets is discussed in chapter 3.

In addition to Congressional Data Sheets, DOD provides two other major system reports to congressional committees--Research, Development, Test and Evaluation Descriptive Summaries and Selected Acquisition Reports. All three reports contain performance information on major weapon systems. However, Congressional Data Sheets focus more specifically on OT&E results than the other reports.

DOD organization responsible for certain OT&E reporting requirements

In 1983, DOD created the Office of the Director for OT&E, pursuant to Public Law 98-94. The law states that the Director will be the principal advisor to the Secretary of Defense on OT&E and the principal official on OT&E matters within the senior management of DOD. The law requires that the Director accomplish such things as (1) analyzing the results of OT&E conducted for each major defense acquisition program and reporting to the Committees on Armed Services and on Appropriations before a final decision to proceed beyond low-rate initial production and (2) preparing an annual report summarizing DOD's OT&E activities during the preceding fiscal year. The Director's responsibilities also include ensuring that DOD's reports on OT&E for major defense acquisition programs are adequate and avoid unnecessary duplication.

At the time production began for the weapon programs included in our review, this office had not been established and therefore, was not involved in the decision to begin production for these weapon programs. In April 1985, a Director was appointed by the President and confirmed by the Senate.

Our reports dealing with OT&E information provided to the Congress

We have been concerned for some time with the quality of test and evaluation information provided to the Congress. In 1979 and again in 1980, we issued reports[5] that stressed the

---

[5]Need for More Accurate Weapon System Test Results to Be Reported to the Congress (GAO/PSAD-79-46, Mar. 9, 1979) and DOD Information Provided to the Congress on Major Weapon Systems Could Be More Complete and Useful (GAO/PSAD-80-24, May 9, 1980).

need for more accurate, complete, and useful information to be included in Congressional Data Sheets. In both reports, we cited instances where conclusions concerning systems limitations were not reported, specific test results and conclusions were omitted, and system weaknesses were not identified.

We performed a limited review of 14 selected Data Sheets for use in hearings on operational testing conducted by the Senate Committee on Governmental Affairs in June 1983. This limited review corroborated many of our previous concerns with Data Sheets.

OBJECTIVES, SCOPE, AND METHODOLOGY

To determine the effect that concurrency has on DOD's ability to predict weapon performance in combat before production start-up, we focused on the provision made to perform OT&E before production start-up, the implementation of the planned OT&E, and the results achieved from OT&E that was performed.

We reviewed the five major weapon systems listed below with the objective of determining whether DOD obtains sufficient OT&E results before production begins on major weapon systems that are concurrently developed and produced. These systems were selected because (1) they were being concurrently developed and produced, (2) certain congressional interest existed with these systems, and (3) they represent major acquisitions from all three military departments.

Air Force	Army	Navy
Air Launch Cruise Missile (ALCM)	Sergeant York air defense gun	F/A-18 aircraft
B-1B bomber		AGM-88A High Speed Antiradiation Missile (HARM)

We discussed the objectives of our review and the weapon choices with representatives of the Office of the Secretary of Defense (OSD), the Office of the DOD Inspector General, and each military department. These officials confirmed that the systems were being concurrently developed and produced and that our sample would enable us to accomplish our objectives.

After selecting the weapon systems, we compared test plans and schedules with actual tests and reviewed test reports to ascertain the risk identified by the testing.

In addition, we reviewed the quality of OT&E information reported in fiscal year 1985 Congressional Data Sheets for eight weapon systems. In selecting the Data Sheets, we chose major weapon systems from all three military departments.

Data Sheets Reviewed		
Air Force	Army	Navy
ALCM	Sergeant York air defense gun	AEGIS cruiser
B-1B bomber	AH-64 advanced attack helicopter	F/A-18 aircraft
Imaging Infrared Maverick Missile		TOMAHAWK Cruise Missile

Our objectives in this segment of the review were to determine

--whether the OT&E results contained in test reports were adequately reflected in the Data Sheets submitted to the Congress,

--the military departments' procedures and practices for preparing the Data Sheets, and

--whether Data Sheets were updated to incorporate the most recent available test information.

We visited several locations during the review to obtain data, views, and ideas relevant to the objectives. The most important locations visited were:

--OSD, Washington, D.C.

--Headquarters Army, Navy, and Air Force, Washington, D.C.

--Various weapon systems program offices.

--U.S. Army Operational Test and Evaluation Agency (OTEA), Falls Church, Virginia.

--U.S. Air Force Operational Test and Evaluation Center (AFOTEC), Albuquerque, New Mexico

--U.S. Navy Operational Test and Evaluation Force (OPTEVFOR), Norfolk, Virginia.

Our review, conducted from September 1983 to July 1984, was performed in accordance with generally accepted government auditing standards. A description of the weapon systems included in our review is in appendix I.

CHAPTER 2

CONCURRENT DEVELOPMENT AND PRODUCTION WITHOUT SUFFICIENT OT&E LIMITS ABILITY TO PREDICT WEAPON PERFORMANCE

Concurrency can be an effective technique to expedite the development and production of weapon systems, provided the practice is well planned and controlled. This requires that adequate safeguards be built into a program to minimize the risk of concurrency. At the very least, these safeguards should provide for performance of at least one phase of OT&E and the completion of planned OT&E before production. Accordingly, the degree of concurrency should remain a planned part of the program and not dictated by uncontrolled or unplanned events. Among other things, this means that delays in scheduled tests must usually be accompanied by corresponding delays in production-decision milestones, and possibly to initial operational capability.[1]

Risks associated with concurrency should be identified and assessed throughout the weapon programs to avoid unplanned delays in scheduled OT&E before production start-up. Adequate OT&E results are important because they provide early identification of problems and can help prevent costly retrofits and performance shortcomings that have occurred with many weapon systems in the past. Furthermore, lack of sufficient and timely test results denies the Congress valuable information on which to base a decision to appropriate production funds for a major weapon program.

For the five weapon systems we reviewed, the program structure for three of the systems called for full OT&E before production. However, as the concurrent weapon programs evolved, the planned OT&E critical to assessing mission performance before production start-up was only partially completed despite DOD's policy on concurrency requiring special attention to OT&E so that performance risks resulting from a shortened acquisition cycle would not affect the planned deployment date. This situation is illustrated in the following table.

---

[1]Initial operational capability--the first attainment of the capability to effectively employ a weapon, item of equipment, or system.

FIVE CONCURRENT WEAPON PROGRAMS

	NAVY		AIR FORCE		ARMY
	F/A-18	HARM	ALCM	B-1B	Sergeant York
Concurrency part of program structure	X	X	●	X	X
Program structure called for full OT&E before production	X	X	X	●	●
Implementation of planned OT&E before production start-up	●	●	●	●	●

X = yes
● = no

This chapter assesses the military departments' use of concurrent development and production in selected major weapon system programs. It also addresses the importance of having sufficient and timely OT&E data available before production start-up.

SLIPPAGE OF SCHEDULED TESTS WITHOUT CONCOMITANT DELAYS IN NAVY PRODUCTION PROGRAMS

Concurrent development and production have long been part of the Navy's F/A-18 and HARM program structure. In the early years of these two programs, the risk of concurrency was to be reduced by scheduling operational tests before key decision points. As the programs evolved, however, slippage in scheduled test and evaluation occurred without concomitant adjustments in program milestones. Both systems were approved for production start-up before test and evaluation in a combat representative environment was complete.

F/A-18 program

The F/A-18 was concurrently developed and produced. The program structure called for completion of some operational testing before production. However, some aspects of the planned OT&E were not completed before starting limited production because of the system's immature development.

In 1979, we expressed concern[2] about concurrency in the F/A-18 program that called for production of 24 to 39 aircraft before a complete Navy operational evaluation.[3] We pointed out that (1) significant risk existed because the aircraft was designed with a new engine, new radar, and new airframe and (2) the tight test schedule allowed little time to correct and retest performance deficiencies that might occur. The program schedule prescribed initial operational testing of the system's potential effectiveness before beginning production. The testing that was envisioned, however, was not accomplished.

The Navy's independent operational testing organization, OPTEVFOR, performed an abbreviated phase of the planned OT&E in March 1980 in support of a limited production decision. However, OPTEVFOR notified the Chief of Naval Operations that it was unable to make a valid assessment of the F/A-18 because of the system's immature development. The initial operational evaluation of the system's potential was not started until October 1980. The Secretary of Defense nonetheless approved limited production of 25 aircraft before OT&E was completed to OPTEVFOR's satisfaction. As shown by the following table, production decisions on the F/A-18 were made before full test results were available.

---

[2]Need to Demonstrate F-18 Naval Strike Fighter Weapon System Effectiveness Before Large-Scale Production (GAO/PSAD-79-25, Feb. 27, 1979).

[3]Operational Evaluation (OPEVAL) is the Navy's operational test to demonstrate operational effectiveness and operational suitability, as well as verification of fixes for problems discovered in development testing. The OPEVAL is conducted using such things as production representative hardware and validated software. The OPEVAL is a term unique to the Navy.

Dates of Test Reports and
Production Decisions on the F/A-18

Planned and actual dates of reports on OT&E results	Planned and actual dates of Production decisions
Initial:	Limited:
Planned: Feb. 1980 Actual: June 1981	Apr. 1980 Apr. 1980
Operational evaluation (fighter):	Full (fighter):
Planned: Oct. 1980 Actual: July 1983	Nov. 1980 June 1981
Operational evaluation (attack):	Full (attack):
Planned: Dec. 1981 Actual: July 1983	Apr. 1982 Mar. 1983

In June 1981 the Secretary of Defense authorized full production of the fighter version. At that time, the F/A-18 had not been operationally tested aboard an aircraft carrier despite OPTEVFOR's warning in 1977 that a

> "major procurement decision for a carrier-based aircraft should not be made without having operationally tested it at sea aboard a carrier."

OPTEVFOR predicted in 1976, that "the program would move ahead whether proper testing had been accomplished or not," and in 1977, that there would be inadequate testing "upon which to base recommendations and decisions at scheduled milestones." The Navy's program schedule specified that the full operational evaluation of the F/A-18's effectiveness and suitability would be completed before the full-production decision but it was not completed until more than 2 years after authorization of the full production of the fighter version. In March 1983, 4 months before the operational evaluation was completed, the Secretary of Defense authorized full production of the attack version.

In commenting on our report, DOD agreed that limited production was approved before OT&E was completed but pointed out that there were encouraging developmental test results that in

DOD's view reduced the risk to an acceptable level. DOD added that the full-production decisions for the fighter version were delayed until June 1981 so that OT&E results from testing during the period of October 1980 to February 1981 could be considered.

While there were Navy operational test reports available when the respective full-production decisions were made, these reports did not provide thorough OT&E results called for in the F/A-18 program schedule and prescribed by DOD and Navy instructions. When full production of the fighter was authorized in June 1981, OPTEVFOR reported on an initial phase of testing intended only to provide an assessment of the aircraft's potential, not a full demonstration of its capabilities in a realistic environment. This type of relatively low level assessment is usually done in support of a decision to begin full-scale engineering development, not full production.

By the time full production was authorized for the attack version, OPTEVFOR had issued a preliminary report on the results of its testing. The preliminary test report did not provide the thorough evaluation of system performance which DOD policy states is needed to support major production commitments.

The full effect of concurrent development and production without sufficient OT&E on the F/A-18 program is difficult to identify. However, certain observations can be made, namely, that expensive retrofits on production models were required to correct problems identified during the operational testing performed after the production commitment was made. The fiscal year 1984 budget for the F/A-18 included $28.5 million for these and other retrofits, plus other modifications and improvements.

HARM program

The pattern of concurrency in the HARM program is similar to the F/A-18. Specifically, (1) concurrent development and production was greater than planned, (2) the initial production decision was made before significant OT&E results were available, and (3) the full-production decision was made before sufficient operational testing was completed.

The HARM program structure indicated that initial operational testing would begin before limited production, but this did not occur. At the time the decision was made to begin limited production, OPTEVFOR had monitored developmental tests and participated in missile firings. However, OPTEVFOR's efforts did not represent testing of the HARM in a realistic operational environment prescribed by DOD guidelines as a prerequisite to production. The following table compares the dates that OT&E results were available with the dates when production decisions were made.

Dates of Test Reports and
Production Decisions for the HARM

Planned and actual dates of reports on OT&E results	Planned and actual dates of Production decisions
Initial:	Limited:
Planned: Jan. 1980 Actual: July 1981	Sept. 1979 Dec. 1980
Operational evaluation:	Full:
Planned: July 1981 Actual: June 1983	Sept. 1981 Mar. 1983

The Navy's fiscal year 1980 budget submission contained a request for limited production of 80 missiles. However, the House Committee on Appropriations refused this request due to

> ". . . insufficient test results, excessive concurrency, and the belief that the desire to enter production arose from adherence to an artificial [initial operational capability] date rather than a rational decision that readiness for production was based on hard evidence."

The Committee concluded that costly retrofitting would result if the missiles had been produced.

The following year the Committee appropriated funds for limited production of the 80 missiles, but also expressed strong concern over the system's poor performance in test firings. The Committee stressed the need for further testing and added a provision that limited production funds not be obligated or spent until the Secretary of Defense certified that, based on test results, HARM was ready for production.

In commenting on our report, DOD agreed that limited production was authorized before operational testing but stated that the results of development testing "with 13 successes out of 18 firings warranted the acceptance of some risk . . ." Ideally, production start-up should be supported by thorough OT&E, particularly if the weapon is being concurrently developed and produced. In December 1980 the Secretary of Defense certified to the Congress that HARM was ready for production although OPTEVFOR's role at that point had been largely limited to monitoring developmental test firings.

The program schedule for HARM called for an operational evaluation of the missile to be completed before authorization of full production, but this did not occur. In commenting on our report, DOD stated that the full-production decision followed completion of operational testing. Our work showed that in March 1983, the project manager requested continuing limited production, and stated that additional operational test results would be presented later to support the full production authorization. The Secretary authorized full production of HARM the following month without the additional operational test results.

Navy officials consider HARM an urgently needed system. Officials in the HARM project office told us that the Navy has, nevertheless, attempted to minimize the risks of concurrency by keeping the production rate low.

THE ARMY'S SERGEANT YORK WEAPON PROGRAM BEGAN PRODUCTION WITHOUT THE BENEFIT OF SUFFICIENT OT&E RESULTS

Concurrent development and production was planned and incorporated into the program structure for the Sergeant York. However, DOD designed the Sergeant York program to have only limited OT&E results available prior to the limited production decision--because the Army judged both performance and cost risks to be acceptable, based on such things as the use of mature components and subsystems and contractor responsibility and experience. Although many components and subsystems have been proven, this is not a substitute for OT&E on the total system before production--particularly the need to ensure that proven subsystems interface properly in the operational environment as part of the total system. With only limited test results, production began on an unproven major system.

The acquisition strategy for the Sergeant York gave the contractor the responsibility to design and produce an air defense gun to meet general Army performance requirements. Throughout the development, the Army adopted a so-called "hands-off" policy. The Army expected this acquisition strategy to result in a system fielded in less time and with less cost growth than in weapon systems acquired along more traditional lines. The Army believed that the Sergeant York was particularly suitable for this type of strategy because it incorporated many off-the-shelf proven subsystems.

In addition, the Army's acquisition strategy included certain provisions in the contract which held the contractor responsible for various corrective actions that may have to be performed. We did not evaluate the adequacy of these contract provisions.

We have issued three reports on the Sergeant York.[4] In two of these reports, we concluded that production should not proceed until sufficient testing of the system's reliability and supportability occurred. DOD's response was that the Sergeant York met most of its performance specifications and was ready for production. DOD acknowledged shortcomings in the system's logistical supportability and that risks existed in proceeding into production, but said these were outweighed by (1) evidence that remedial actions had been identified, (2) the urgency to field a new air defense gun, and (3) cost savings achievable by not delaying production.

In 1983 we issued the third report which stated:

> "The Army's plans to test the reliability and maintainability of its new Sergeant York air defense gun had to be abandoned when the prototype the prime contractor delivered for testing in May 1982 was found to be unacceptable."

In commenting on our 1983 report, DOD officials said that the decision to terminate the government tests stemmed from several factors. These included (1) the related realization that certain prototype subsystems were close to wearing out after 2-1/2 years of constant testing and (2) the prototype's lack of a number of features to be included in the production model.

In 1983 our concern was that there would be little authenticated performance information to measure the reliability and maintainability of the Sergeant York until 2 years after production began. We recommended that the Army assess the Sergeant York's progress before deciding to exercise the second production option.

DOD did not agree to formally require an assessment of Sergeant York's performance before exercising the second production option. DOD officials said Army assessments are made continuously. We believed that it was important to have a formalized assessment to establish accountability, particularly since

---

[4] Inherent Risk in the Army's Acquisition Strategy Demands Particular Caution in Evaluating the Division Air Defense Gun System's Production Readiness (C-PSAD-80-9, Jan. 31, 1980).

Tests and Evaluations Still in Progress Should Indicate Division Air Defense Gun's Potential Effectiveness (C-MASAD-82-7, Feb. 26, 1982).

The Army Should Confirm Sergeant York Air Defense Gun's Reliability and Maintainability Before Exercising Next Production Option (MASAD-83-8, Jan. 27, 1983).

Sergeant York had less government testing and evaluation than usual.

We testified on the Sergeant York, along with the DOD Inspector General, before the Senate Committee on Governmental Affairs on September 28, 1984. The DOD Inspector General stated that the Sergeant York was tested in late 1981, several months before the production decision. This test had two elements--a developmental test portion and an operational test portion. The Inspector General stated that the Army did not release the complete OT&E results until after the May 1982 production decision. Furthermore, the Inspector General faulted the production decisionmaking process for the Sergeant York because of questionable representations of developmental test results by the Army--including oversimplified and selective analysis of development test and overly optimistic performance assumptions.

In our testimony, we explained that the Sergeant York's potential in combat had not been fully demonstrated. In 1984, another limited operational test[5] was performed that demonstrated certain performance shortcomings. We also stated that the Army had several additional tests planned, including a test to ensure that the system's reliability is improved. As part of a separate ongoing review, we are continuing to monitor the system's performance in a combat representative environment.

In commenting on this report, DOD agreed that the Army undertook an unconventional acquisition strategy for the Sergeant York program. DOD stated that preliminary test data from a "short check test (combined DT/OT)" indicated that the weapon program had the potential to be effective, and therefore sufficient data was available to support going ahead with production. However, deficiencies and shortcomings found during the "short check test" were to be corrected and verified during a subsequent test. The subsequent test did not demonstrate that all deficiencies and shortcomings were corrected because the Army was constrained by too few prototypes, no spare parts, nor enough time to correct the deficiencies and shortcomings that surfaced in the short check test.

THE AIR FORCE'S ALCM AND B-1B WEAPON PROGRAMS BEGAN PRODUCTION BEFORE OT&E RESULTS WERE AVAILABLE

In both Air Force programs included in our review (ALCM and the B-1B bomber), sufficient OT&E data did not exist before production start-up because the Air Force was committed to prescribed time schedules rather than demonstrating performance capabilities before production.

---

[5]This test was characterized as a limited operational test because certain operational testing elements were missing, such as trained troops and support equipment and spares.

ALCM

The Air Force had limited and inconclusive OT&E results available for its preproduction evaluation of ALCM. AFOTEC, an independent test agency, issued a report before the production decision stating that:

> ". . . Operational effectiveness test results are inconclusive . . . Several unplanned events precluded an adequate evaluation of the most important performance and mission planning capabilities. These events included four missile crashes, inadequate test range instrumentation performance, and late delivery of mission planning components. The planned operational test was of limited scope due to program constraints, e.g., limited number of missiles (none of which were production configured) . . . Consequently, the unplanned events seriously impacted the generation of statistically meaningful quantities of data."

AFOTEC was concerned that conclusions drawn from operational testing might not be accurate because testing was not operationally realistic.

We also reported[6] before the initial production decision, that operational testing presented an incomplete assessment of aircraft equipment. We stated that many questions remained about the system's performance and recommended that the Secretary of Defense initially approve only a limited production quantity of ALCMs, postponing the full-scale production decision until remaining tests could be done to confirm the total system's capabilities.

DOD placed a high priority on ALCM deployment to preclude shortfalls in strategic weapons in the 1980s. Accordingly, a highly concurrent schedule was established despite the problems discussed above. DOD approved the missile for production before all planned OT&E results were available. In a later report,[7] we stated that the production decision may have been prudent and expedient, but if the problems were not quickly resolved, the system might be deployed with severe operational limitations.

AFOTEC did follow-on operational testing for 3 years after the April 1980 ALCM production decision. There were constraints in the testing, including (1) a shortage of prescribed test

---

[6]Cruise Missiles: Status and Issues As They Near Production (GAO/C-PSAD-80-19, Feb. 28, 1980).

[7]Some Land Attack Cruise Missiles Acquisition Programs Need to Be Slowed Down (GAO/C-MASAD-81-9, Feb. 28, 1981).

time and (2) test range restrictions which prevented accomplishing completely representative OT&E. The system nevertheless proceeded through production and fixed program milestones with limited, compressed testing and persistent deficiencies.

After the Air Force began ALCM deployment, system deficiencies identified through testing persisted. For example:

--The lack of adequate complete subsystems affected the Mission Planning System, which is key to ALCM effectiveness. The prototype system, however, was not built to meet real-world requirements, and it could not produce mission data adequate to support realistic testing. Despite the continuing evolution of software, the system remained problematic through 1983.

--The navigational accuracy of ALCM was questioned after testing in 1979 and remained an issue into 1983. The inaccuracy could result in missile crashes if not corrected.

--Range realism has been a continuing concern to the Air Force; the combination of ranges used for flight testing adversely affected the credibility of test results. AFOTEC managers reported that the terrain did not resemble the expected operational environment, nor did the ranges provide enough continuous territory to test the missile as it was intended to be employed. Finally, in 1984--2 years after deployment, the United States gained access to air space that would allow for more realistic ALCM testing.

In commenting on our report, DOD stated that the ALCM underwent 10 combined DT/OT flights before the production decision and that fixes to identified deficiencies were "in-hand and the risk of putting the system into production was acceptable." DOD also stated that deficiencies identified were corrected and verified in subsequent launches. We did not evaluate the results of the most recent ALCM launches because the test report had not been issued at the time of our review. However, some of the problems with the ALCM existed for several years and could have been identified through effective OT&E before a decision was made to begin production.

B-1B bomber

A high degree of planned concurrency is being employed in the B-1B program. As discussed previously, DOD's policy requires that special attention be given to having OT&E results available before production; however, this risk reducing safeguard was not adopted in the B-1B program.

DOD issued a study on the B-1B the same day that President Reagan announced the production decision--October 2, 1981. The study included analyses of risks in several program areas, such as:

--Schedule risks:

> "The proposed concurrent development and production schedules are success oriented and increase the risk that acquisition of logistics elements such as peculiar support equipment, spare and repair parts, technical manuals, maintenance training, and training equipment could be delayed due to late availability of stable engineering design data. Delays could hinder the Air Force's ability to develop an organic support capability by the Initial Operational Capability date . . . .

--Logistics assessment:

> "The overall risk associated with being able to acquire and field a supportable B-1 Variant weapon system in accordance with the proposed schedule is judged to be moderate to high. The predominant risk is attributed to . . . concurrency of the full-scale development and production programs . . . .
>
> "The risk associated with acquiring . . . supply support prior to initial deployment of the B-1 Variant is assessed to be moderate to high . . . spare and repair parts ordered against the required delivery schedules to support operational deployment may be rendered obsolete by engineering changes during the flight test program thus requiring modification or replacement . . . ."

In April 1983, we reported[8] that time available for flight testing before the initial operational capability date was limited, especially for evaluating new defensive avionics. Our report noted that flight tests showed that the defensive avionics countermeasures system had not matured "to a level needed for operational testing."

The B-1B did not begin OT&E until 3 years after the October 1981 production decision. AFOTEC personnel documented program challenges to B-1B OT&E, which include integration of avionics. They noted that even though individual subsystems may have operated successfully in other environments, the possibility exists that their integration in the B-1B could cause interface problems. Interoperability of avionics should be tested and evaluated before the initial operational capability.

---

[8]The B-1 Bomber Program--A New Start (GAO/MASAD-83-21, Apr. 13, 1983).

Air Force officials are optimistic about the B-1B program. They contend that the high priority and visibility of the program fosters rapid response to program problems, and that they will meet the operational capability date with a proven, effective system.

In September 1984,[9] we issued a report that focused on the planned logistics support for the B-1B bomber. We said that the inadequacy of the logistics data developed during research and development of the B-1B's predecessor--the B-1A--and the concurrent development and production schedule necessitated by a congressionally mandated fielding date, forced Air Force planners to make logistics support decisions before they had sufficient data to support them. This increased the risk that operating and support costs could be more than they may have been had normal DOD development procedures been employed before starting production.

In commenting on a draft of this report, DOD stated that the risks of concurrency were weighed against the overriding need "expressed by the President and directed by the Congress to immediately embark upon the nation's strategic modernization." As suggested by DOD in its comments, this report has been modified to recognize a DOD study which assessed the risk surrounding the production decision for the B-1B program. The DOD study indicated that the B-1B program was assessed as having moderate to high risk in several program areas. The predominant risk was attributed to the concurrent development and production of the program.

CONCLUSIONS

Production began for all five concurrent weapon systems programs we reviewed without accomplishing the OT&E that had been planned in order to reasonably assure that the system would meet operational requirements. Sufficient and timely OT&E is particularly essential in a concurrent program because of the greater risks that exist in producing a weapon system while development is still underway.

Without conducting planned OT&E critical to assessing mission performance, decisionmakers do not have important information needed in deciding whether or not to advance systems to production. This increases the likelihood that weapons will be produced without sufficient knowledge about their performance in a representative operational environment.

---

[9]Logistics Support Costs for the B-1B Aircraft Can Be Reduced (GAO/NSIAD-84-36, Sept. 20, 1984).

RECOMMENDATIONS

We recommend that the Secretary of Defense, in an effort to improve the quality of information to decisionmakers in the acquisition of major weapon systems, should assure through the Office of the Director of OT&E that test results critical to assessing mission performance are available before production start-up.

We also recommend that the Secretary of Defense require the Office of the Director of OT&E to provide its views to the Senate and House Committees on Appropriations and Armed Services on the impact of not performing OT&E critical to assessing mission performance before production begins. This is particularly important when planned OT&E is not accomplished.

AGENCY COMMENTS AND OUR EVALUATION

DOD stated that the programs we reviewed did not complete as much OT&E as is usually required before a weapon system is authorized to proceed into production, but other factors such as urgency of the requirement and the cost of delay, persuaded DOD decisionmakers to authorize proceeding into production. DOD further stated that in almost all cases, the production rate was limited until OT&E was completed and test results could be considered before authorizing full production. However, DOD agreed with our recommendation on improving the use of concurrency and stated that the Director of OT&E will be able to assure that adequate operational testing is planned to support decision milestones established for future major weapon systems acquisitions.

In the draft of this report, we proposed that when sufficient OT&E results are not available, the Congress should be given justification by the Director of OT&E concerning the necessity and risks of advancing concurrent systems into production.

DOD disagreed with this proposal stating that its decisionmakers must decide when the risk of advancing a system into production is acceptable. DOD stated that the Director will make recommendations known to the Secretary of Defense and the Chairman of the Defense System Acquisition Review Council but the final decision will be based on many factors, such as the urgency of the operational requirement and the avoidance of cost increases.

We agree that the final decision for advancing a system to production rests with the Secretary of Defense and we therefore

modified our recommendation to call for the establishment of a procedure requiring the Office of the Director of OT&E to provide its views to the Congress. Our fundamental objective in making this recommendation is to assure that congressional decisionmakers are fully informed of the impact of starting production when sufficient OT&E results are not available.

CHAPTER 3

THE QUALITY OF CONGRESSIONAL DATA SHEETS NEED FURTHER IMPROVEMENT TO BE MORE USEFUL TO THE CONGRESS

OT&E results are contained in test reports that are often lengthy and technically complex. Congressional Data Sheets provide a means of condensing these reports into concise, useful summaries. Properly prepared, the Data Sheets reveal not only the known strengths and weaknesses of weapon systems, but also the limitations of testing and the required operating capabilities yet to be demonstrated.

We reviewed the OT&E results reported in fiscal year 1985 Congressional Data Sheets for eight major weapon systems. These Data Sheets were submitted to the Congress in early calendar year 1984. We found that in all cases, the Data Sheets omitted or did not accurately portray important test and evaluation results or related information which we believe would have been useful to the Congress in assessing weapon system performance and the production fund request. For example, our review revealed that:

--The Data Sheet on the Navy's AEGIS class cruiser did not discuss certain problems that were cited in test reports, such as the need to improve computer program performance which is important for the system to be operationally effective.

--The Air Force's ALCM Data Sheet did not accurately portray the weapon system's performance in a combat representative environment. Certain tests where the missiles crashed were described as a partial success because other objectives were accomplished.

Test results omitted or not accurately portrayed can create a misleading impression of weapon systems' demonstrated performance; thus denying the Congress complete and accurate information to make budgetary decisions on major weapon system programs.

PREPARING AND REVIEWING OT&E SECTIONS OF CONGRESSIONAL DATA SHEETS

OSD issues guidance to the military services on the preparation of annual Congressional Data Sheets. We found that the military instructions have replicated without elaborating OSD's general guidance. The OSD guidance requires that the Data Sheets

--describe and summarize the information on recently completed, ongoing, and future planned testing;

--identify and evaluate the effect of differences between the article tested and the article to be procured;

--identify any important subsystems not tested and estimate the probable effect on test results; and

--identify important discrepancies found and corrective action taken or planned.

Procedures for preparing Data Sheets differs among the military departments

The military services independent test agencies[1] are an important resource in preparing Data Sheets. As requested by the Chairman, we tracked the process of reporting operational test data from the services' independent testing organizations up through the chain of command within DOD. The independent operational test agencies in the Navy and the Air Force prepare the OT&E sections of the Data Sheets. In the Army, however, many organizations are responsible for test and evaluation and for providing information for the OT&E section of the Data Sheets--which contributed to some of the problems with the Army's Data Sheets identified later in this chapter.

Navy procedures for preparing the Congressional Data Sheets

The Navy's independent test agency, OPTEVFOR, prepares the OT&E sections of the Data Sheets. Navy instructions stress the importance of accurate, timely Data Sheets and provide for several levels of review. Still, as discussed later in this chapter, the documents excluded certain information we believe would be useful in the congressional review process. Most of the problems we found in the Navy's Data Sheets occurred in OPTEVFOR's original drafts of the OT&E sections, and remained unchanged throughout all levels of review.

Pursuant to the instruction from the Office of the Chief of Naval Operations, OPTEVFOR prepares initial drafts of the OT&E sections of Data Sheets. OPTEVFOR then sends its drafts of the test sections to the Chief of Naval Operations.

The Chief of Naval Operation's Office of Research, Development, Test and Evaluation reviews the drafts for currency, accuracy, and composition and sends them to OSD and the Navy's

---

[1]In this report, we refer to the military departments operational test agencies as being independent because they do not report to the military organization responsible for developing or using the weapon system.

program coordinators for review. It also sends draft copies to OPTEVFOR to verify that the OT&E sections remain accurate and up to date. These drafts are returned to the Naval Operations Office and revised if necessary. The final drafts are collated with the budget data, and sent to the Navy's comptroller to be submitted to the Congress.

Air Force procedures for preparing Congressional Data Sheets

The Air Force's procedures are similar to the Navy's. The Air Force's budget directorate lists programs requiring test and evaluation data. AFOTEC, the Air Force's independent test agency, prepares most of the initial drafts of the OT&E sections of the Data Sheets and submits them to Air Force Headquarters.

Most of the problems that we found with the Air Force's Data Sheets appeared in the initial drafts prepared by AFOTEC. The drafts generally remained unchanged through the DOD review process.

Army procedures for preparing Congressional Data Sheets

Unlike the Navy and Air Force, the Army's independent test agency does not prepare the initial drafts of Army Data Sheets. The Army's Office of the Deputy Chief of Staff for Research, Development, and Acquisition is responsible for the test and evaluation sections of the Data Sheets. Various Army test organizations provide information to be included. The Deputy Chief sends the test and evaluation section to the program manager, the Army's OTEA, and OSD for review and comment. The Deputy Chief also consolidates the test and evaluation section with the remaining sections of the Data Sheets.

We recently reported[2] that the involvement of the many Army groups in test and evaluation on weapon systems resulted in a complex organizational structure and in piecemeal presentation and reporting of test and evaluation results. Considerable effort was being spent by the Army in accumulating and interpreting test data. The fact that each organization looked only at a portion of the total test results hindered a complete portrayal of the collective effect of deficiencies identified during testing.

After our study, the Army devised a pilot program to provide a comprehensive evaluation of selected weapon systems in varying stages of development. OTEA is integrating the work of the many

---

[2]The Army Needs More Comprehensive Evaluations to Make Effective Use of Its Weapon System Testing (GAO/NSIAD-84-40, Feb. 24, 1984).

analysis and test groups. It will consolidate evaluations of each system's progress in developmental and operational effectiveness and will continue to review the test and evaluation section of the Congressional Data Sheets.

DATA SHEETS FOR SELECTED NAVY SYSTEMS OMIT IMPORTANT TEST RESULTS AND INFORMATION ON TESTING LIMITATIONS

The three Navy Data Sheets we reviewed omitted information on test results which created misleading impressions of weapon systems' demonstrated performance. Furthermore, a misleading impression was also created because limitations in the test and evaluation process were not made evident, as illustrated in the Data Sheet on the AEGIS cruiser.

AEGIS cruiser

OT&E of the AEGIS weapon system began in 1977, but the system was not tested with the Ticonderoga, the first of the AEGIS/CG-47 class cruisers, until 1982. In April 1983, OPTEVFOR "in the most intensive short-of-war operational testing of a surface ship in the history of our Navy," tested the weapon system and the ship in simulated antiair, antisurface, and antisubmarine warfare scenarios. The tests involved numerous engagements against both live and simulated targets and included actual missile launches by the AEGIS weapon system.

OPTEVFOR published the results of these tests in June 1983, yet the fiscal year 1985 Data Sheet included neither detailed results of the AEGIS missile firings nor OPTEVFOR's assessment of key subsystems performance. The Data Sheet also omitted an important conclusion reported by OPTEVFOR. During an earlier phase of testing in 1981, OPTEVFOR found that although AEGIS had the potential to be operationally effective and suitable, the system had limitations and problems, such as the need to improve computer program performance. OPTEVFOR concluded that although the AEGIS weapon system and the CG-47 remained potentially effective and suitable, the limitations and problems found in 1981 had not been completely resolved. However, the Data Sheet cited only those conclusions on the system's potential. It omitted OPTEVFOR's conclusion that known limitations and problems remained unresolved.

As a part of earlier antiair warfare tests in 1981, OPTEVFOR assessed AEGIS' probability of success in various operational environments. The various types of environments were cited in OPTEVFOR's June 1981 report, but only the results of the less severe types of environments were included in the Data Sheet. Also, the Data Sheet strongly implied that one of AEGIS' antiair capabilities was demonstrated during operational testing, when in fact, the capability was demonstrated only during earlier developmental testing.

F/A-18 strike fighter

In July 1983, OPTEVFOR completed its operational evaluation of the F/A-18. OPTEVFOR's report included assessments of the aircraft's effectiveness, suitability, and readiness for full fleet introduction, but expressed strong concern about limitations to the operational evaluation.

The fiscal year 1985 Congressional Data Sheet listed test limitations cited in the report and noted that some of them precluded a complete "evaluation of operational effectiveness and operational suitability in the strike warfare mission . . . " However, the Data Sheet gave no indication of the significance or effect of the incomplete evaluation. In contrast, OPTEVFOR's report stated that the effect of these limitations left unresolved a number of critical issues. The Data Sheet also omitted OPTEVFOR's conclusion that of

> "greater significance [than the test limitations]. . . was the lack of several definitive required operational capabilities and performance thresholds, particularly for the F/A-18's attack configuration . . . While this lack of specificity did not affect the structure and conduct of operational testing . . . it did affect [OPTEVFOR's] ability to evaluate test results."

A similar lack of perspective was reflected in the way the Data Sheet addressed OPTEVFOR's comparison of F/A-18 and A-7 aircraft capabilities and the test agency's assessment of the F/A-18's aircraft carrier launch and recovery requirements. The Data Sheet quoted OPTEVFOR's conclusions and recommendations on these issues but did not include brief passages from the test report illustrating the effect that deployment of the F/A-18 will have on battle groups and individual carriers.

Furthermore, the Data Sheet on the F/A-18 did not mention the follow-on OT&E results available 4 months before the Data Sheet was submitted to the Congress. The OT&E report noted that the earlier operational evaluation did not adequately establish the F/A-18's range and payload capabilities as a strike fighter, partly because the strike fighter role had not been formally defined. The test report addressed the aircraft's capability to execute strike missions from carriers and marine airfields, and noted that deficiencies found during the earlier evaluation still existed during these follow-on tests. The Navy did not incorporate these or other parts of the test report into the Data Sheet.

TOMAHAWK cruise missile

The fiscal year 1985 Data Sheet on the TOMAHAWK noted that initial operational testing was in process but did not mention

the results cited in two interim OPTEVFOR reports. These reports, based on preliminary and incomplete analysis of results from tests of the 1981-83 period, concluded that both versions of the TOMAHAWK were potentially effective and suitable. However, the reports recommended that the missiles be approved for limited fleet introduction only if corrections were made.

These test reports also gave detailed information on TOMAHAWK test firings during three segments of operational testing in 1981-83. The Data Sheet addressed neither the scope nor the results of these test firings, nor did it report that problems with the TOMAHAWK caused the Chief of Naval Operations to suspend operational testing in 1982 and again in 1983.

POTENTIAL EFFECTS OF KNOWN DEFICIENCIES OMITTED FROM THREE AIR FORCE DATA SHEETS

In comparing the OT&E results included in test reports to Air Force Data Sheets, we found that significant issues were omitted. For example, the Data Sheets did not contain all significant information on

--test limitations and their effects on test results and knowledge about the weapon system and

--weapon system deficiencies (including problems with operational effectiveness, supportability, and maintainability) and the known and potential effects of those deficiencies.

ALCM

OT&E results on ALCM provided valuable information on the weapon systems performance in a combat representative environment. However, the Data Sheet described certain launched missiles that crashed as a partial success because other objectives were accomplished. The project office determined that accomplishment of any of the numerous mission objectives results in a partial success. However, DOD officials stated that another Air Force organization, the Strategic Air Command, views such launches as failures because all mission objectives were not met.

The Data Sheet indicated major problems with suitability, and identified computer software deficiencies involving subsystems such as the diagnostic Electronic System Test Set. The Data Sheet did not indicate, however, that these deficiencies cause low levels of confidence in some subsystems' performance and require extensive manpower to work around shortcomings, as indicated in the test report.

According to DOD officials, the ALCM manpower problem was not included in the Data Sheet because the problem was scheduled

to be resolved. One official explained that some important test information was omitted from the Data Sheet because major milestones had been accomplished and items were being produced and deployed. We believe that the Air Force should have described the ALCM manpower problem in the Data Sheet and then updated the document when the problem was corrected.

B-1B bomber

Test reports have not been issued yet for the B-1B, therefore, we could not compare test results to the Data Sheet. AFOTEC, however, documented potential problems with the B-1B, such as the integration and testing of offensive and defensive avionics.

Air Force officials said that detailed information on the avionics subsystems was not included because the potential degrading effects of integrating the subsystems are speculative, and it was considered inappropriate to identify a problem not yet fully proven.

We believe the information on the avionics should have been included in the Data Sheets. The B-1B is a derivative of the B-1A, which underwent approximately 1,900 hours of flight test from 1974 to 1981. DOD uses test data from the B-1A aircraft in assessing the B-1B program. The B-1A program showed that the defensive avionics subsystem had not been sufficiently developed to adequately evaluate system effectiveness. Because the integration of avionics could also degrade mission performance, this information is critical in assessing the B-1B program and should have been included in the Data Sheet.

Imaging Infrared Maverick Missile

The Data Sheet for the Imaging Infrared Maverick Missile identified the critical issues and deficiencies discovered during OT&E. However, it excluded conclusions relating to effects on operational suitability and effectiveness.

The Data Sheet noted continuing concerns about the weapon's reliability. It did not, however, include the problem that was identified in the test report with dormant reliability (i.e., performance after storage). The test report said that dormant reliability remains a potential problem and it must be tested as soon as possible.

Air Force officials stated that the Maverick program is so highly visible that the Air Force has always provided critical test information to the Congress immediately, rather than via the Data Sheet. Nonetheless, we believe the Air Force should also include important information on OT&E results in Data Sheets as required by law. We believe that conclusions about the effect of

deficiencies should be included in Data Sheets even if they were reported through other means.

DATA SHEETS FOR SELECTED ARMY SYSTEMS DO NOT ALWAYS EXPLAIN SIGNIFICANCE OF TEST LIMITATIONS

The two Army Data Sheets we reviewed omitted certain information on testing limitations and safety related deficiencies that should have been included.

Sergeant York

The Data Sheet on the Sergeant York did not explain the significance of reduced testing. For example, to demonstrate the durability and reliability characteristics of the Sergeant York, which is mounted on a tank chassis, a test was planned to fire 15,000 rounds of ammunition and to travel 4,000 miles. Because of time limits, only about 3,600 rounds were fired and, due to frequent subsystem failures, the Sergeant York traveled less than 300 miles. The Data Sheet did not indicate the planned testing to be the minimum required to address specific critical issues. Rather, the Data Sheet noted only that the number of events made were less than planned.

Test results indicated that computer and software problems were encountered during testing of the fire control subsystem. The primary power unit experienced occasional shutdown and hydraulic failures. Under low temperature conditions, the ammunition feed system in its tested configuration jammed and dented ammunition. In a combat situation, failure to meet any of these requirements would either cause a mission to be terminated, an engagement to be aborted, or performance to be degraded below required performance levels. The Data Sheet reported that the reliability of the tested system declined during the test and attributed it to excessive wear on the system and insufficient spare parts to support the tests.

The objective of the Sergeant York's maintenance concept, of which built-in test equipment is an important part, is to keep the system in a high state of readiness without an enormous investment in spare parts and maintenance personnel. Built-in test equipment is supposed to identify failures in the system's performance and pinpoint the malfunctions to specific subsystems. This is accomplished by sensors which relay the failure to the operator or gunner's display panel. For example, if the turbine overheats, the sensor activates an alerting lamp on the display panel to alert the crew. The test report showed that the built-in test equipment was incapable of isolating faults to the responsible subsystem primarily because the built-in test equipment was not fully developed. Although the Data Sheet mentioned

that the test was designed to demonstrate built-in test capability, it did not report that the testing organization considered the equipment inadequate.

The Data Sheet did not identify any of the safety related deficiencies reported by the test agency, nor did it identify any corrective action taken or planned by the Army to resolve these deficiencies. For example, the Data Sheet should have disclosed that the arm that operates the switch to allow the driver to move between the crew compartment and turret was found bent on two occasions. With the switch inoperative, anyone or anything caught between the turret and the stationary chassis could be sheared.

We recognize that the Army instructions for preparing Data Sheets do not specifically require that safety related deficiencies must be reported. However, these deficiencies were identified through test and evaluation and may significantly affect operational effectiveness and suitability after deployment.

AH-64 helicopter

The configuration of the helicopter engine used during operational testing was not the configuration to be procured. The production engine--which is similar to the tested engine--was subsequently flight tested. However, the Data Sheet did not identify and report the effect of the differences between the engine tested and the engine to be procured.

We reported[3] that the operational test consisted of threats of the 1970s. Simulated Soviet air defense weapons of the late 1980s, which the AH-64 will face were not used.

CONCLUSIONS

Congressional Data Sheets do not always provide the Congress with an adequate description and assessment of the available information on OT&E results of major weapon systems. In 1979, we reported[4] that if Data Sheets are to be effective, they must contain a complete, objective, and current presentation of weapon system capabilities and risks, as demonstrated through operational testing. We believe that this is still not being done because neither the military departments nor OSD are ensuring that the Data Sheets contain accurate and complete OT&E information.

---

[3]The Army Needs More Comprehensive Evaluations to Make Effective Use of Its Weapon System Testing (GAO/NSIAD-84-40, Feb. 24, 1984).

[4]Need for More Accurate Weapon System Test Results to Be Reported to the Congress (GAO/PSAD-79-46, Mar. 9, 1979).

In 1980, we reported[5] that DOD made some improvements to its Data Sheets. For example, at that time, we found that DOD was doing an excellent job of reporting the conclusions and recommendations resulting from OT&E. However, for the major weapon systems selected in our current review, this situation did not exist.

DOD's general policy guidance for preparing Congressional Data Sheets is adequate. However, the military services' implementing instructions have only replicated OSD's general policy on preparing Data Sheets. The services' implementing instructions should more explicitly state the importance of these documents and the specific type of information that should be included in the Data Sheets. Our review showed that OT&E results reported to the Congress do not consistently disclose key aspects of major weapon system testing, such as a system's

--acquisition risk;

--critical issues;

--operational effectiveness and suitability;

--reliability, availability, and maintainability; and

--test environment.

RECOMMENDATION

To improve the reporting of OT&E results to the Congress, we recommend that the Secretary of Defense direct the military departments to expand their implementing instructions on preparing Congressional Data Sheets to include a more complete and accurate portrayal of OT&E results on major weapon programs.

AGENCY COMMENTS AND OUR EVALUATION

DOD agreed that Congressional Data Sheets have not provided accurate or complete OT&E information to the Congress. DOD stated that in the future the Office of the Director of OT&E will be responsible (1) for ensuring that the Data Sheets contain accurate and complete OT&E information and (2) for clearer reporting of operational test results.

DOD concurred with our recommendation and stated that the Office of the Director of OT&E will review the existing instructions on preparing Data Sheets and then revise the instructions within 6 months, if necessary.

---

[5]DOD Information Provided to the Congress on Major Weapon Systems Could be More Complete and Useful (GAO/PSAD-80-24, May 9, 1980).

DESCRIPTION OF SELECTED WEAPON SYSTEMS

AIR FORCE SYSTEMS:

ALCM

ALCM is designed to be a long-range, subsonic, jetpowered vehicle armed with a nuclear warhead which uses sophisticated navigational aids for flying at low altitudes, avoiding detection, and striking targets with a high degree of accuracy. The Air Force estimates that it will cost $4.6 billion to purchase 1,787 missiles.

B-1B bomber

The B-1B program revives the nuclear and conventional warfare-capable manned bomber that was curtailed by President Carter's cancellation of B-1A production and deployment in 1977. The Air Force estimates that it will cost $28.3 billion to purchase 100 bombers. Improvements in the aircraft should enable the B-1B to penetrate high-threat environments at low levels and high speed.

Imaging Infrared Maverick Missile

The Imaging Infrared Maverick Missile is designed to be a precision guided munition improved for night and adverse weather use, depending on new imaging infrared guidance for its enhanced capabilities. The Air Force estimates that it will cost $5.8 billion to purchase 60,697 missiles.

ARMY SYSTEMS:

AH-64 Advanced Attack Helicopter

The AH-64 is an antiarmor helicopter. The AH-64 should be able to destroy enemy ground targets, including tanks, from the air. This weapon system should be able to operate at night and in adverse weather conditions. The Army estimates that it will cost $7.3 billion to purchase 524 helicopters.

Sergeant York Air Defense Gun

Sergeant York, formerly known as the Division Air Defense Gun, is being developed to fill a perceived air defense void in the forward battle area. The Sergeant York should be able to engage armed helicopters, fixed-winged aircraft, lightly armored vehicles, trucks, and personnel. The Army estimates that it will cost $4.2 billion to purchase 622 Air Defense Guns.

NAVY SYSTEMS:

AEGIS cruiser

The AEGIS/CG-47 cruiser is designated for assignment to carrier battle groups or surface action groups, and should provide quick reaction defense against aircraft and antiship missiles. The Navy estimates that it will cost $28.8 billion to purchase 26 cruisers.

AGM-88A HARM

The HARM is a guided missile system designed to help tactical aircraft penetrate enemy defenses by destroying or suppressing the radar used to direct surface-to-air missiles and antiaircraft guns. The joint Navy and Air Force HARM program will provide missiles for use on the F/A-18 and other Navy and Air Force planes. The Navy plans to purchase about 8,100 missiles for an estimated $3.3 billion. As of March 31, 1984, the Navy had received 103 production missiles and the Air Force had received 21.

F/A-18 aircraft

The F/A-18 naval strike fighter is a twin-engine, single-pilot aircraft designed to replace the F-4 and the A-7 and perform fighter and attack missions for the Navy and the Marine Corps. The Navy plans to purchase 1,377 F/A-18s at an estimated cost of $40 billion. The Navy plans to begin deploying F/A-18s aboard carriers in 1985.

TOMAHAWK cruise missile

The TOMAHAWK weapon system is designed to deliver long-range nuclear or conventional cruise missiles against land and sea targets. Two Navy versions are (1) a nuclear armed, submarine-launched missile for use against land targets and (2) a conventionally armed, submarine-launched antiship missile. The Navy estimates that it will cost $13 billion to purchase 4,068 missiles.

OFFICE OF THE SECRETARY OF DEFENSE
WASHINGTON, D.C. 20301

12 FEB 1985

OPERATIONAL TEST AND EVALUATION

Mr. Frank C. Conahan
Director, National Security and
International Affairs Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Conahan:

This is the Department of Defense (DoD) response to your letter of December 26, 1984, which transmitted your draft report (GAO Code 951820) entitled "Production of Some Major Weapon Systems Began with Only Limited Operational Test and Evaluation Results" (OSD Case No. 6663).

The GAO report correctly concluded that an acquisition program with planned concurrency should provide special attention to operational test and evaluation (OT&E) so that performance risks resulting from a shortened acquisition time do not affect the planned deployment date. The new office of Operational Test and Evaluation in OSD will ensure that OT&E planning for concurrent acquisition programs is given more attention than it has received in the past.

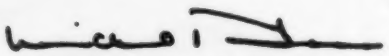
The acquisition programs reviewed in the report did not complete as much OT&E as is usually required before a weapon system is authorized to proceed into production but other factors, such as urgency of the requirement and the cost of delay, persuaded DoD decision makers to authorize proceeding into production.

In almost all cases, the production rate was limited until OT&E was completed and test results could be considered before authorizing full production. For some programs, the availability of OT&E results was difficult to recognize because decision events preceded issuance of the OT&E final report. However, OT&E briefings and interim reports were used to present operational test results to decision makers in support of most of the full production decisions.

It is also agreed that the quality of Congressional Data Sheets can be improved. OSD and the Services are committed to this effort. In addition, the OT&E Annual Report from the Office of the Director of Operational Test and Evaluation will supplement the OT&E information in the Congressional Data Sheets.

DoD comments addressing each of the findings and recommendations contained in the draft report are in the Enclosure to this letter. The opportunity to comment on the report in draft form is appreciated.

Sincerely,



MICHAEL D. HALL
Brigadier General, USAF
Office of the Director
Operational Test and Evaluation

Enclosure

GAO DRAFT REPORT - DATED DECEMBER 26, 1984
(GAO CODE 951820 - OSD CASE NO. 6663)

"PRODUCTION OF SOME MAJOR WEAPON SYSTEMS BEGAN WITH ONLY LIMITED OPERATIONAL TEST AND EVALUATION RESULTS"

DEPARTMENT OF DEFENSE COMMENTS

* * * * *

FINDINGS

- FINDING A: DoD Acquisition Policy Permits Concurrency Between Development And Production In Order To Minimize Acquisition Time. GAO reported that DoD policy on system acquisition stresses the importance of minimizing acquisition time and, to accomplish this objective, permits the Services to build concurrency into acquisition programs. GAO reported that, under DoD policy, concurrency may be justified if risks have been fully identified, assessed and accepted in a planned acquisition strategy, but that concurrency is unjustified if it evolves in an uncontrolled manner after the strategy is established, or as a result of testing delay or other unplanned events. GAO found, however, that DoD policy provides that the degree of concurrency should be based on savings in acquisition time balanced against the cost, risk and mission urgency of the acquisition program. GAO concluded that an acquisition program with planned concurrency should provide special attention to operational test and evaluation (OT&E) so that performance risks resulting from a shortened acquisition time do not affect the planned deployment date. (p. 2, GAO Draft Report)

  DoD Concurs. The new office of the Director of Operational Test and Evaluation will be reviewing acquisition programs with planned concurrency and ensuring that plans for operational testing are structured to support decision milestones for the program.

- FINDING B: Meeting Test And Evaluation Objectives Is Key Element Of Acquisition Decisions. GAO reported DoD policy on operational test and evaluation provides that accomplishment of test and evaluation objectives is to be a key requirement to support the level of commitment at each key decision point. GAO concluded that production startup, an important milestone, must be supported by sufficient OT&E to estimate operational efffectiveness and suitability, including logistic supportability. (p. 5, GAO Draft Report)

  DoD Partially Concurs. DoD concurs that accomplishment of test and evaluation objectives is a key requirement to support the level of commitment at each key decision point; however, limited production start-up need only be supported by sufficient test and evaluation to insure that

requirements have been met or necessary fixes are well enough understood to be included in early production equipment. The necessary test and evaluation may be development testing, operational testing or a combination of both. Confirmation of operational effectiveness and suitability may be delayed until production equipment is available to be used as test articles.

- FINDING C: Insufficient OT&E Completed Before Initial Production Decision And Before Final Production Decisions For Both Fighter And Attack Versions Of Navy's F/A-18 Program. GAO reported that the F/A-18 was concurrently developed and produced. The program structure called for completion of some operational testing before production. GAO found, however, that sufficient OT&E was not completed before starting production because of the system's immature development. GAO noted that in 1979, it had expressed concern about concurrency in the F/A-18 program. At that time GAO had concluded significant risk existed because the aircraft was designed with a new engine, new radar and new airframe and the tight test schedule allowed little time to correct and retest any performance deficiencies that might occur. The GAO also pointed out that although the Navy's independent operational testing organization (OPTEVFOR) performed a short phase of OT&E in March 1980, it notified Navy officials it was unable to make a valid assessment because of the F/A-18 system immature development. GAO reported that nonetheless in April 1980 limited production of 25 aircraft was approved. Subsequently, in June 1981, full production of the fighter version was authorized before its OT&E was completed and before the aircraft had been operationally tested aboard a carrier. Again, in March 1983, full production of the attack version of the F/A-18 was authorized. While noting the full effect of concurrent development without sufficient OT&E in the F/A-18 program is difficult to measure, GAO concluded that certain observations can be made--i.e., that expensive retrofits to production models of the F/A-18 were required to correct problems indentified during the operational testing conducted after the production commitment was made. (p. 13-15, GAO Draft Report)

  DoD Partially Concurs. Although limited production was approved prior to completion of OT&E to OPTEVFOR's satisfaction, there were encouraging development testing results that in the view of the decision makers reduced the risk to an acceptable level. The full production decision for the fighter version was delayed until June 1981 in order that OT&E results from testing during the period of October 1980 to February 1981 could be considered. Full production of the attack version wasn't authorized until March 1983 after operational testing from May 1982 to October 1982 which included eight days of CV operations on board USS Constellation.

- **FINDING D: Concurrence in HARM Program.** GAO found that the HARM program experience to be similar to what it found on the F/A-18--i.e., (1) concurrent development and production was greater than planned, (2) the initial production decision was made before significant OT/E results were available, and (3) the full-production decision was made before the operational evaluation was completed. GAO reported that the HARM program schedule indicated that initial operational testing would begin before limited productions, but this did not occur. GAO noted the concern of Congress over the HARM system's poor performance and its requirement that the Secretary certify that HARM was ready for production. GAO concluded that although in December 1980 it was certified to the Congress that HARM was ready for production, in fact, at that time only developmental tests had been performed. (p. 16-17, GAO Draft Report)

  DoD Partially Concurs. Although limited production of HARM missiles was authorized in November 1980 before operational testing, the results of development testing with 13 successes out of 18 firings warranted the acceptance of some risk in order to avoid cost increases and to satisfy an urgent Navy operational need. The full production decision in March 1983 followed completion of operational testing by the Navy (NOV 81 -NOV 82) and operational testing by the Air Force (NOV 81 - OCT 82). The certification to the Congress that HARM was ready for production in December 1980 was based on a balance of technical risk, cost impact, and urgent fleet need and was not questioned by the Congress.

*- **FINDING E: Army's AH-64 Was Made A Concurrent Program, And Production Was Delayed Until Evaluation Of Test Results Became Available.** The GAO found that, whereas concurrency was not planned at the inception of the AH-64 program, a decision to concurrently produce the system was made after judging the technical risks, cost increases and schedule delays that would otherwise occur. GAO concluded that because insufficient time was allowed to obtain adequate OT&E results, the schedule for the production decision was delayed. (pp. 18-20, GAO Draft Report)

  DoD Non-Concurs. A decision was not made to concurrently produce the AH-64. In fact, the production decision in April 1982 followed after consideration of results from combined DT/OT testing that was completed in August 1981. The initial production awarded in April 1982, was additionally limited to eleven aircraft and further procurements held contingent upon OSD review of further test data on the target acquisition designation sight (TADS) subsystem.

- **FINDING F: Despite Concurrency, The Sergeant York Program Had Only Limited And Selective Test Results As The Basis For A Production Decision.** Because the Army judged risks to be

*An earlier draft of this report contained a discussion on the AH-64 advanced attack helicopter. We have deleted the discussion of the AH-64 based on DOD's comment that a decision was not made to concurrently develop and produce it.

acceptable (i.e., based on mature components and subsystems and an experienced contractor), GAO found that DoD decided to accept both concurrency and reliance on a combined developmental and operational test before production. The GAO concluded, however, that the use of proven subsystems does not relieve the need for OT&E to assure that these interface properly in an operational environment. The GAO also found that not all deficiencies were corrected during testing because the Army was constrained by having too few prototypes, no spare parts, and not enough time to correct deficiencies. The GAO noted that in September 1984 the DoD Inspector General testified that although a test was conducted in late 1981, several months before the production decision, the Army did not release the results until after the May 1982 production decision. The production decision was, instead, based upon a selective analysis of developmental test results which showed overly optimistic performance assumptions. (p. 21, p. 24, GAO Draft Report)

DoD Partially Concurs. The Army undertook an unconventional acquistion strategy for the SGT YORK program. Competitive testing by two contractors limited the amount of government testing possible. Preliminary test data results in January 1982 from a short check test (combined DT/OT) indicated that SGT YORK had the potential to be effective. In October 1984, during Senate hearings on SGT YORK, the Principal Deputy for the Undersecretary of Defense for Research and Engineering testified that sufficient data was presented during the DSARC III process to support going ahead with production. Recently the Secretary of Defense delayed ordering additional systems until specific operational testing is conducted to verify SGT YORK's effectiveness.

- FINDING G: Air Force Had Inconclusive OT&E Data For Preproduction Evaluation Of Air Launched Cruise Missile. GAO reported that the Air Force, according to its own test report, had only limited and inconclusive OT&E on which to base the preproduction evaluation of the Air Launched Cruise Missile (ALCM). Because DoD placed a high priority on deployment of the ACLM to preclude shortfalls in strategic weapons in the 1980s, GAO found that a highly concurrent schedule was established. Despite identified problems, the missile was approved for production before planned OT&E results were available. GAO concluded that as a result, after deployment began in 1981, system deficiencies indentified through testing persisted. (The GAO cited as examples of such deficiencies lack of adequate subsystems and the level of navigational accuracy.) (pp. 25-28, GAO Draft Report)

DoD Partially Concurs. The ALCM underwent ten combined DT&E/IOT&E flights prior to the production decision. Both the April 1980 AFSARC and SDARC concluded that fixes to identified deficiencies were in-hand and the risk of putting the system into production was acceptable. Time has proven that to have been a prudent decision. Between June 1980 and October 1982, twenty more combined DT&E/FOT&E flights were

conducted. Of these, fourteen were completely successful and three were partially successful. Deficiencies have not "persisted." The GAO cited "adequate subsystems and navigational accuracies" is understood to be in reference to the SAC mission planning system and its interface with the missile. All deficiencies identified were corrected and subsequently verified in the aforementioned FOT&E launches.

- FINDING 4: B-1B Aircraft Began Production Three Years Before OT&E Scheduled To Begin. GAO reported that the B-1B aircraft was not scheduled to begin OT&E until three years after the October 1981 production decision. The reason was that the Air Force was committed to a prescribed time schedule rather than to providing sufficient resources to demonstrate before proceeding into production. GAO concluded that not only did DoD eliminate the preproduction OT&E called for by its policy, it also failed to assess adequately the increased risk of producing a concurrent weapon system without OT&E results. (p. 25, pp. 28-30, GAO Draft Report)

  DoD Non-Concurs. The DoD did adequately assess the risk of concurrent development and production of the B-1B. More importantly, the risks of concurrency were weighed against the overriding need, expressed by the President and directed by the Congress to immediately embark upon the nation's strategic modernization. The decision process weighed all factors, including OT&E adequacy and, the risks of concurrency were determined to be acceptable.

- FINDING I: Congressional Data Sheets On Major Systems Do Not Provide Accurate Or Complete OT&E Information. Based on its review of eight systems, GAO found that the 1985 Congressional Data Sheets either omitted or did not accurately portray important test and evaluation information that would have been useful to the Congress in assessing production funding requests. GAO concluded that OT&E results reported to the Congress do not consistently disclose such key aspects of system test and evaluation as acquisition risk; critical issues; operational effectiveness and suitability; reliability, availability and maintainability; and the test environment. According to GAO, if the Data Sheets are to be effective, they must contain a complete, objective and current presentation of system capabilities and risks, as demonstrated through operational testing. GAO concluded that this is still not being done because neither the Military Departments nor OSD is ensuring that the Data Sheets contain accurate and complete OT&E information. (p. 33, pp. 46-47, GAO Draft Report)

  DoD Concurs. One of the responsibilities of the new office of the Director of Operational Test and Evaluation in OSD is to ensure that Congressional Data Sheets contain accurate and complete OT&E information. The Office of the Director of Defense Test and Evaluation and the Services are also committed to improving the quality of future Congressional Data Sheets.

- FINDING J: Procedures For Preparing Congressional Data Sheets On Major Systems Differ Among Military Departments. GAO reported that procedures differ among the Military Departments on the preparation of Congressional Data Sheets on major systems, principally in terms of the role of the Services independent test agencies. GAO found that in the Navy and Air Force, the operational test agencies prepared the OT&E sections of the Data Sheets. By contrast, GAO reported, many organizations are responsible for performing OT&E and providing OT&E information for the Army Data Sheets. GAO concluded that this diffusion of responsibility contributed to problems it identified in Army Data Sheets. Moreover, while the AH-64 was operationally tested with an engine different from the one to be procured, GAO found that the Data Sheet failed to report either the differences between the two versions or the durability of the production configuration. GAO further concluded that the Services' implementing instructions for the Data Sheets should more explicitly state the importance of the Data Sheets as well as the specific information which should be included. (pp. 35-37, 47, GAO Draft Report)

  DoD Concurs. The Office of the Director of Operational Test and Evaluation in OSD will review the existing instruction and then revise any, if necessary, within six months.

- FINDING K: Congressional Data Sheets On Three Navy Systems Omitted Important Test Information. GAO found that Congressional Data Sheets on three Navy Systems (AEGIS, F/A-18 and TOMAHAWK) omitted information on test results, thereby creating misleading impressions about demonstrated system performance. For example, the Data Sheet on the F/A-18 failed to mention the results of follow-on OT&E, which indicated that deficiencies found in earlier testing still existed. These test results were available four months before submission of the Data Sheet. GAO also noted that, in the case of the AEGIS cruiser, the Data Sheet was misleading in not making clear limitations in the test and evaluation process. (The Data Sheet implied that the anti-air capabilities of the system were demonstrated in operational testing, while this capability had actually been demonstrated only in earlier, developmental testing.) (pp. 37-40, GAO Draft Report)

  DoD Concurs. Review by the office of the Director of Operational Test and Evaluation in the future will stress clearer reporting of operational test results.

- FINDING L: Significant Issues Omitted From Three Air Force Data Sheets. GAO found that a comparison of OT&E reports with Data Sheets for three Air Force systems showed that significant issues were omitted. GAO reported the Data Sheets in question (on ALCM, B-1B and Maverick) did not contain all significant information either on the effects of testing limitations on knowledge of system performance, or on weapon system deficiencies and their known or potential

effects. Examples cited by GAO were the description of certain ALCM launches as partial successes even through the missiles crashed, and the failure of the B-1B Data Sheet to mention potential avionics problems which had been documented by the Air Force testing agency. (pp. 41-43, GAO Draft Report)

DoD Concurs. Review by the office of the Director of Operational Test and Evaluation in the future will stress clearer reporting of operational test results.

- **FINDING M: Data Sheets On Two Army Systems Omitted Information On Testing Limitations And Safety Related Deficiencies.** GAO reported that its review of Data Sheets on two Army Systems (Sergeant York and AH-64) showed that information had been omitted on testing limitations and safety related deficiencies. While testing of Sergeant York was much less than planned in terms of number of rounds fired and travel time, for example, this limitation was never explained in the Data Sheet. According to GAO, that document also failed to mention any of the safety related deficiencies identified by the Army test agency, or the corrective actions taken or planned to resolve them.

  DoD Concurs. Review by the office of the Director of Operational Test and Evaluation in the future will stress clearer reporting of operational test results.

RECOMMENDATIONS

- **RECOMMENDATION 1:** GAO recommended that the Secretary of Defense, through the Office of the Director, OT&E, should assure, sufficient test results are available to warrant production startup. (p. 32, GAO Draft Report)

  DoD Concurs. The office of the Director of Operational Test and Evaluation through review and approval of Test and Evaluation Master Plans, review and approval of Service operational test plans, participation in both the DSARC decision process and the Defense Resources Board decision process, as a principal member, will be able to assure that adequate operational testing is planned to support decision milestones established for future major weapon system acquisitions.

  RECOMMENDATION 2: GAO recommended that when sufficient OT&E results are not available, the Congress should be given justification by the Director, OT&E, concerning the necessity and risks of advancing systems into production. (p. 32, GAO Draft Report)

  DoD Non-Concurs. Decision makers in DoD must decide when the risk of advancing a system into production is acceptable. The Director of Operational Test and Evaluation will make his recommendations known to the Secretary of

Defense and the Chairman of the DSARC but the final decision will be based on many factors, such as: the urgency of the operational requirement, the avoidance of cost increases and the results of developmental testing, in addition to operational testing results. The Director will submit a report to the Congress before a final decision is made within DoD to advance a system beyond Low Rate Initial Production (LRIP).

RECOMMENDATION 3. GAO recommended that the Secretary of Defense direct the Military Departments to expand their implementing instructions on preparing Congressional Data Sheets to include a more complete and accurate portrayal of OT&E results on major weapon programs. (p. 47, GAO Draft Report)

DoD Concurs. The office of the Director of Operational Test and Evaluation will review the existing instructions and then revise any, if necessary, within six months.

(951820)